Conditional generative adversarial networks to generate pseudo low monoenergetic CT image from a single-tube voltage CT scanner

Yoshinori Funama^{a,1}, Seitaro Oda^b, Masafumi Kidoh^b, Yasunori Nagayama^b, Makoto Goto^c, Daisuke Sakabe^c, Takeshi Nakaura^b

^a Department of Medical Radiation Sciences, Faculty of Life Sciences, Kumamoto University, Kumamoto, Japan

^b Department of Diagnostic Radiology, Faculty of Life Sciences, Kumamoto University, Kumamoto, Japan

^c Department of Radiology, Kumamoto University Hospital, Kumamoto, Japan

ARTICLE INFO

Keywords:

Virtual monochromatic image
Conventional tube-voltage image
Dual-energy computed tomography
Conditional generative adversarial networks

ABSTRACT

Purpose: To generate pseudo low monoenergetic CT images of the abdomen from 120-kVp CT images with cGAN. **Materials and Methods:** We retrospectively included 48 patients who underwent contrast-enhanced abdominal CT using dual-energy CT. We reconstructed paired data sets of 120 kVp CT images and virtual low monoenergetic (55-keV) CT images. cGAN was prepared to generate pseudo 55-keV CT images from 120-kVp CT images. The pseudo 55 keV CT images in epoch 10, 50, 100, and 500 were compared to the 55 keV images generated using peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM). **Results:** The PSNRs were 28.0, 28.5, 28.6, and 28.8 at epochs 10, 50, 100, and 500, respectively. The SSIM was approximately constant from epochs 50 to 500. **Conclusion:** Pseudo low monoenergetic abdominal CT images were generated from 120-kVp CT images using cGAN, and the images had good quality similar to that of monochromatic images obtained with DECT software.

1. Introduction

Dual-energy computed tomography (DECT) can improve material differentiation using two different X-ray energy spectra [1,2]. This procedure has been used in important clinical applications, such as virtual monochromatic imaging, iodine mapping, and assessment of effective atomic number and electron density [2–4]. Low-energy virtual monochromatic imaging is used to obtain high iodine contrast between lesions and adjacent tissues compared with single polychromatic 120-kVp image. Moreover, it is highly useful in clinical settings [5–7]. Uyeda et al. [8] have reported that virtual monochromatic imaging with low keV has increased the conspicuity of noncalcified gallstones, thereby improving their detectability. Moreover, Lv et al. have reported that the image quality of virtual monochromatic images (VMIs) at 55 keV and 62 keV using DECT were similar to that of 80-kVp and 100-kVp images, and VMIs at 50 keV had higher overall image quality [9]. However, unlike the most common types of CT scans that can yield single-tube voltage images, DECT can be obtained using specific technologies, such as the dual-layer spectral detector CT, kilovoltage switching, and dual-source CT [2].

In 2014, Goodfellow et al. [10] have introduced generative adversarial networks (GANs), which are one of the recent interesting topics in computer science. GANs are generative models, and with the use of training data, the new data obtained were similar to the training images generated using GANs. Conditional GAN (cGAN) is a type of GAN that can control a generated image under conditions determined in advance [11]. Recently, cGANs were found to have state-of-the-art performance in several image generation tasks, including text-to-image synthesis, super-resolution, and image-to-image translation [12–15].

VMIs are applied using only DECT scanners in limited healthcare facilities [16]. If VMIs are obtained from single polychromatic 120-kVp images using cGAN, DECT scanners will no longer be required, and this can be achieved in various facilities. In addition, unlike VMIs obtained using DECT scanners, those generated with cGAN are obtained using post-processing function after single-tube voltage scanning if necessary. In our study, we focused on abdominal contrast-enhanced VMIs with low energy and the feasibility of cGAN in clinical settings.

Thus, the current study aimed at generating pseudo low monoenergetic CT images of the abdomen from standard single-tube voltage CT images with cGAN. Moreover, the feasibility of obtaining pseudo low

E-mail address: funama@kumamoto-u.ac.jp (Y. Funama).

¹ Department of Medical Radiation Sciences, Faculty of Life Sciences, Kumamoto University. 4-24-1 Kuhonji, Kumamoto 862-0976, Japan.

monoenergetic CT images with cGAN was assessed.

2. Material and methods

The local institutional review board approved the protocol of this retrospective study, and the need for informed consent was waived.

2.1. Patients

The study cohort included 51 adult patients with renal dysfunction (estimated glomerular filtration rate <45 mL/min/1.73 m²) who underwent abdominal CT scan with 50% reduced iodinated contrast media between December 2016 and March 2017. Among these patients, three whose injection speed was reduced using a 20-G catheter were excluded from the study. The remaining 48 patients were enrolled (30 men and 18 women; mean age: 74.0 ± 9.7 [range: 45–89] years; mean weight: 58.5 ± 10.4 [range: 40.0–76.0] kg).

2.2. CT scanning protocols and image reconstruction

All patients underwent dual-layer DECT using the IQon Spectral CT System (Philips Healthcare, Cleveland, OH, the USA). DECT can obtain VMIs at different monochromatic X-ray energies (keV) under the standard tube-voltage setting. The single-tube voltage for obtaining VMIs was set at 120 kVp. The following scanning parameters were used for all image acquisitions: detector configuration, 64×0.625 mm (detector collimation); gantry rotation time, 0.5 s; and helical pitch (beam pitch), 0.8. Tube current modulation with automatic exposure control was set using an image quality reference (dose right index [DRI]; Philips Healthcare). The DRI was based on the relationship between the reference diameter and the water-equivalent diameter, which was calculated according to the cross-sectional attenuation characteristics of the patients by the CT scanner. In this study, CT scanning was performed with a DRI of 22 to maintain constant image quality regardless of patient attenuation characteristics.

To generate pseudo low monoenergetic CT image from 120-kVp CT image with cGANs, we reconstructed paired data sets of standard tube voltage (120 kVp) CT images and virtual low monoenergetic (55 keV VMIs) CT images. Thirty examinations (6744 images) were used as training data and 18 examinations (3954 images) as test data. The training data were used to train the generator and discriminator networks and to determine the degree of learning by cGAN. The test data set was used only after the final training. All images were acquired at a slice thicknesses of 1.0 mm and slice interval of 1.0 mm, with an abdominal standard kernel (C) in a 30–35-cm display field of view based on a patient's body size. For image reconstruction, hybrid iterative reconstruction (IR) with IR level 3 (iDose⁴ level 3; Philips Healthcare) was applied to 120-kVp images and 55-keV VMIs to reduce image noise [17,18]. The IR levels are used to define the strength of the IR technique in reducing image quantum mottle noise. The image noise is decreased with increasing IR levels under the same radiation dose level. The images with IR level 3 is available for theoretical 22.5% noise reduction compared with filtered back projection images [19,20]. A medical workstation software (IntelliSpace Portal; Philips Medical) was used to generate paired data sets of 120-kVp and 55-keV CT images.

2.3. Model architecture and training

To generate pseudo low monoenergetic CT image, we used pix2pix, a modified conditional GAN using Keras 2.1.5 (<https://keras.io>) [21,22] and TensorFlow 1.5.0 (Python libraries) [23,24] and performed the parameter adjustments for monochrome images. Similar to other GAN models, pix2pix has generator and discriminator architectures. The maximum resolution that can be handled with pix2pix is a 256×256 matrix. Therefore, CT images were resized with Python libraries (Pillow 5.1.0) to a 256×256 matrix. The details of the two networks are

presented in Fig. 1. In this study, a generator network was used to generate a paired pseudo low monoenergetic CT image from the 120-kVp CT image (Fig. 1a). For the generator network, we used the U-Net [25], an encoder–decoder network with skip connections between mirrored layers in the encoder and decoder stacks (Fig. 1b). For the discriminator network, a convolutional PatchGAN classifier [24,26], which only penalized the structure at the scale of image patches, was utilized (Fig. 1c). The input of the PatchGAN was either the pair of 120-kVp CT image and 55-keV VMI or the pair of 120-kVp CT image and the pseudo low monoenergetic CT image. This discriminator was used to classify whether an input 256×256 matrix image pair was real (120-kVp CT images and real VMIs) or fake (120-kVp CT images and generated VMIs), to provide a real image pair for a 30×30 matrix of probability, and to obtain the average of all probabilities. Moreover, the inventors of the pix2pix network have recommended this discriminator. The patchGAN outputs 30×30 local real image probabilities and takes the average of these probabilities as the real image probability of the entire images. Both the generator and discriminator used the modules in the form of convolution–BatchNorm–ReLU [27,28]. We applied the Adam solver [29], with a learning rate of 0.0002 and momentum parameters $\beta_1 = 0.5$ and $\beta_2 = 0.999$.

2.4. Quantitative analysis

After training, the pseudo 55-keV CT images generated using test data at epochs 10, 50, 100, and 500 were compared with the 55-keV VMIs. To evaluate the differences between the 55-keV VMIs and the generated pseudo 55-keV CT images, the peak signal-to-noise ratio (PSNR) [30] and structural similarity (SSIM) index were calculated [31]. The PSNR is used to calculate the magnitude of average error between the 55-keV VMIs and the pseudo-55 keV CT images. It is the ratio between the maximum possible value (power) of a signal and the power of distorting noise that affects the quality of its representation, and it is calculated using the following equation:

$$PSNR = 20 \times \log_{10} \left(\frac{MAX_f}{\sqrt{MSE}} \right),$$

Where MAX_f is the maximum possible pixel value that the original image format can take. The MSE was calculated using the following equation:

$$MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (f_{ij} - g_{ij})^2$$

where f and g are the images to be compared, both of size $M \times N$ [30].

The SSIM was used to measure the structural similarity between two images, with value ranging from -1 to 1 . When two images were nearly identical, their SSIM was close to 1 . The PSNR and SSIM were calculated using 128×128 pixels at the center of the 256×256 CT images to calculate these indices with limitations to actual structures because there is no signal at the edge of the CT images.

3. Results

3.1. Quantitative image analysis

Fig. 2 shows the PSNR between the 55-keV VMIs and the generated pseudo 55-keV CT images at varying epochs. The MSE requiring the calculation of PSNR is presented in Table 1. The PSNRs were 28.0 ± 3.6 , 28.5 ± 3.7 , 28.6 ± 3.8 , and 28.8 ± 3.8 at epochs 10, 50, 100, and 500, respectively. Moreover, the PSNR was approximately stable from epochs 50 to 500.

Fig. 3 shows the SSIM between the 55-keV VMIs and the generated pseudo 55-keV CT images at varying epochs. The SSIM were 0.94 ± 0.09 , 0.95 ± 0.09 , 0.95 ± 0.09 , and 0.95 ± 0.09 at epochs 10, 50, 100, and 500, respectively. The SSIM was constant from epochs 50 to 500.

Figs. 4 and 5 show the representative cases for generating pseudo 55-keV CT images in patients with multiple hepatocellular carcinoma and pancreatic cancer liver metastases.

In this study, we generated pseudo low monoenergetic CT scan images of the abdomen from original single-tube voltage CT images using cGAN. Moreover, the feasibility of performance obtaining pseudo low

Fig. 2. Peak signal-to-noise ratio between the 55-keV VMI and the generated pseudo 55-keV CT images at different number of epochs. The body of the boxplot consists of the first quartile, median, and third quartile. The lower and upper whiskers on the vertical line represent the minimum and maximum values.

Table 1

Mean square error (MSE) at different epochs.

Epoch number	MSE	
	Mean	SD
10	191.7	391.7
50	179.4	385.5
100	181.8	396.0
500	179.3	396.2

SD: Standard deviation.

monoenergetic CT images with cGAN was evaluated.

The pseudo 55-keV CT images generated from the 120-kVp CT images using cGAN were similar to the 55-keV VMIs obtained from DECT scan. In our study, the PSNR and SSIM were used as the image metrics. A higher PSNR value indicates a higher image quality. In contrast, a lower PSNR indicates high numerical differences between the 55-keV VMIs and the generated pseudo 55-keV CT images. Moreover, the SSIM is an important image quality metric used to measure the similarity between the 55-keV VMIs and the generated pseudo 55-keV CT images, and it is correlated to the perceived quality of the human visual system [31]. The PSNR and SSIM had almost constant values from epochs 50 to 500, and the pseudo 55-keV CT images had sufficient image quality compared with the actual 55-keV VMIs (SSIM > 0.94).

We believe that the GAN-based architectures are promising techniques that can be used in obtaining medical images from existing ones. Previous studies have performed the following tasks: positron emission tomography (PET)-CT translation, correction of magnetic resonance motion artifacts, PET denoising, metal artifact reduction, and noise reduction from low dose CT [13,15,32,33]. Another study has shown a fully automated segmentation method with high accuracy for the knee cartilage and meniscus [12]. In general, GANs are a type of neural

Fig. 3. The structural similarity (SSIM) index between the 55-keV VMI and the generated pseudo 55-keV computed tomography scan images at different number of epochs. The body of the boxplot consists of the first quartile, median, and third quartile. The lower and upper whiskers on the vertical line represent the minimum and maximum values.

network architecture that can generate new data on their own, and they have gained a lot of attention in the field of computer science. cGAN is an extension of GANs and learn a conditional generative model. An image-to-image translation using GANs had impressive outcomes in terms of image enhancement, super-resolution, and artistic style transfer [13]. In our study, we generated pseudo low monoenergetic CT images of the abdomen from 120-kVp CT images using cGAN (pix2pix network). The pix2pix model is a type of cGAN in which the generation of the output images is conditional based on an input (120-kVp CT images in this case). For the discriminator network, a convolutional PatchGAN classifier was used. The input of the PatchGAN is either the pair of the 120-kVp CT images and 55 keV VMIs or the pair of the 120-kVp CT images and the pseudo low monoenergetic CT images. The discriminator must determine whether the pair of input images (pair of 55-keV VMIs and 120-kVp CT images or pair of pseudo low monoenergetic CT image) is real or fake.

Several studies have shown the efficacy of VMIs [8,9,16]. However, specific types of CT scan require DECTs, and these CT scans are not widely used in clinical settings compared with conventional CT scan. In addition, a CT operator must select the DECT mode or conventional CT mode before the CT scanning except for the dual-layer spectral detector CT. If pseudo VMIs with various monoenergies are obtained using conventional CT scanners in clinical settings, diagnostic improvements are achieved without the use of DECT scanners. In our study, although the pseudo 55-keV CT images were generated from 120-kVp CT images, VMIs with monoenergies 40 and 50 keV can also be obtained using cGAN. The advantage of the pseudo VMIs with various monoenergies is post-processing image creation using cGAN after conventional 120-kVp CT scanning without DECT capabilities.

The present study had several limitations. First, we focused on 55-keV VMIs, which are commonly used in clinical settings, and VMIs with different monoenergies were not assessed. Second, this study included a small number of patients at a single institution. Moreover, to

Fig. 4. Pseudo 55-keV computed tomography (CT) images of an 81-year-old female patient with multiple hepatocellular carcinoma. 120-kVp CT image (a), virtual 55-keV VMI (b), and pseudo 55-keV CT image (epoch 500). (c) In visual inspection by a radiologist, the pseudo 55-keV image is almost similar to the virtual 55-keV image compared with the original 120-kVp image. However, the visual contrast of the pseudo 55-keV image is not exactly the same as that of the virtual 55-keV image.

Fig. 5. Pseudo 55-keV computed tomography (CT) images of a 53-year-old female patient with pancreatic cancer liver metastases. 120-kVp CT image (a), virtual 55-keV VMI (b), and pseudo 55-keV CT image (epoch 500). (c) Although the contrast of the pseudo 55-keV image is not exactly the same as that of the virtual 55-keV image, the contrast between liver tumors and hepatic parenchyma was increased in the pseudo 55-keV image compared to that in the virtual 55-keV image.

train a generalized model, it is necessary to use CT images with different scanning parameters as well as from different machined types and manufacturers. Therefore, further large-scale studies must be conducted to validate our results in the effect of spatial resolution, noise texture, and other factors. Lastly, we used an image resolution of 256×256 because of the limitation of the pix2pix architecture. For this reason, we could not determine how well conditional GANs perform at 512×512 , the normal CT resolution. Next step, we need to compare a 512×512 obtained via a simple interpolation with python with the 512×512 images of the ground truth VMIs.”

In conclusion, pseudo low monoenergetic abdominal CT images were generated from original single-tube voltage CT images using cGAN, and the images had good quality similar to that of monochromatic images. Thus, cGAN can be a promising tool in generating pseudo low-energy CT images from standard-dose abdominal CT images.

References

- [1] Johnson TR. Dual-energy CT: general principles. *AJR Am J Roentgenol* 2012;199: S3–8.
- [2] McCollough CH, Leng S, Yu L, Fletcher JG. Dual- and multi-energy CT: principles, technical approaches, and clinical applications. *Radiology* 2015;276:637–53.
- [3] Goodsitt MM, Christodoulou EG, Larson SC. Accuracies of the synthesized monochromatic CT numbers and effective atomic numbers obtained with a rapid kVp switching dual energy CT scanner. *Med Phys* 2011;38:2222–32.
- [4] Pelgrim GJ, van Hamersvelt RW, Willemink MJ, Schmidt BT, Flohr T, Schilham A, et al. Accuracy of iodine quantification using dual energy CT in latest generation dual source and dual layer CT. *Eur Radiol* 2017;27:3904–12.
- [5] Leng S, Yu L, Fletcher JG, McCollough CH. Maximizing iodine contrast-to-noise ratios in abdominal CT imaging through use of energy domain noise reduction and virtual monoenergetic dual-energy CT. *Radiology* 2015;276:562–70.
- [6] Megibow AJ, Sahani D. Best practice: implementation and use of abdominal dual-energy CT in routine patient care. *AJR Am J Roentgenol* 2012;199:S71–7.
- [7] Yu L, Christner JA, Leng S, Wang J, Fletcher JG, McCollough CH. Virtual monochromatic imaging in dual-source dual-energy CT: radiation dose and image quality. *Med Phys* 2011;38:6371–9.
- [8] Uyeda JW, Richardson LJ, Sodickson AD. Making the invisible visible: improving conspicuity of noncalcified gallstones using dual-energy CT. *Abdom Radiol (NY)* 2017;42:2933–9.
- [9] Lv P, Zhou Z, Liu J, Chai Y, Zhao H, Guo H, et al. Can virtual monochromatic images from dual-energy CT replace low-kVp images for abdominal contrast-enhanced CT in small- and medium-sized patients? *Eur Radiol* 2019;29:2878–89.
- [10] Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial nets. In: *Proceedings of the 27th international conference on neural information processing systems – volume 2*. Montreal, Canada: MIT Press; 2014. p. 2672–80.
- [11] Mirza M, Osindero S. Conditional generative adversarial nets. *Adv Nets* 2014;1411: 784.
- [12] Gaj S, Yang M, Nakamura K, Li X. Automated cartilage and meniscus segmentation of knee MRI with conditional generative adversarial networks. *Magn Reson Med* 2019.
- [13] Yi X, Walia E, Babyn P. Generative adversarial network in medical imaging: a review. *Med Image Anal* 2019;58:101552.
- [14] Wang Y, Zhou L, Wang L, Yu B, Zu C, Lalush DS, et al. Locality adaptive multi-modality GANs for high-quality PET image synthesis. *Med Image Comput Comput Assist Interv* 2018;11070:329–37.
- [15] Yi X, Babyn P. Sharpness-aware low-dose CT denoising using conditional generative adversarial network. *J Digit Imaging* 2018;31:655–69.
- [16] Kang H-J, Lee JM, Lee SM, Yang HK, Kim RH, Nam JG, et al. Value of virtual monochromatic spectral image of dual-layer spectral detector CT with noise reduction algorithm for image quality improvement in obese simulated body phantom. *BMC Med Imaging* 2019;19:76.
- [17] Hou Y, Liu X, Xu S, Guo W, Guo Q. Comparisons of image quality and radiation dose between iterative reconstruction and filtered back projection reconstruction

- algorithms in 256-MDCT coronary angiography. *AJR Am J Roentgenol* 2012;199: 588–94.
- [18] Sakabe D, Funama Y, Taguchi K, Nakaura T, Utsunomiya D, Oda S, et al. Image quality characteristics for virtual monoenergetic images using dual-layer spectral detector CT: comparison with conventional tube-voltage images. *Phys Med* 2018; 49:5–10.
- [19] Noël PB, Fingerle AA, Renger B, Münzel D, Rummeny EJ, Dobritz M. Initial performance characterization of a clinical noise-suppressing reconstruction algorithm for MDCT. *AJR Am J Roentgenol* 2011;197:1404–9.
- [20] Miéville FA, Gudinchet F, Brunelle F, Bochud FO, Verdun FR. Iterative reconstruction methods in two different MDCT scanners: physical metrics and 4-alternative forced-choice detectability experiments—a phantom approach. *Phys Med* 2013;29:99–110.
- [21] Dogaru R, Dogaru I. BCONV – ELM: binary weights convolutional neural network simulator based on Keras/Tensorflow, for low complexity implementations. In: 2019 6th international symposium on electrical and electronics engineering (ISEEE); 2019. p. 1–6.
- [22] Gulli APS. Deep learning with Keras. Packt Publishing Ltd; 2017.
- [23] Rau A, Edwards PJE, Ahmad OF, Riordan P, Janatka M, Lovat LB, et al. Implicit domain adaptation with conditional generative adversarial networks for depth prediction in endoscopy. *Int J Comput Assist Radiol Surg* 2019;14:1167–76.
- [24] Isola P, Zhu J-Y, Zhou T, Efros AA. Image-to-image translation with conditional adversarial networks. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR) (IEEE, 2017); 2017. p. 5967–76.
- [25] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Medical image computing and computer-assisted intervention – MICCAI 2015. Springer; 2015. p. 234–41.
- [26] Klages P, Benslimane I, Riyahi S, Jiang J, Hunt M, Deasy JO, et al. Patch-based generative adversarial neural network models for head and neck MR-only planning. *Med Phys* 2020;47:626–42.
- [27] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. In: Francis B, David B, editors. Proceedings of the 32nd international conference on machine learning. Proceedings of machine learning research: PMLR; 2015. p. 448–56.
- [28] Glorot X, Bordes A, Bengio Y. Deep sparse rectifier neural networks. In: Geoffrey G, David D, Miroslav D, editors. Proceedings of the fourteenth international conference on artificial intelligence and statistics. Proceedings of machine learning research: PMLR; 2011. p. 315–23.
- [29] Kingma D, Adam Ba L. A method for stochastic optimization. *ICLR* 2015. 2015.
- [30] Horé A, Ziou D. Image quality metrics: PSNR vs. SSIM. In: 2010 20th international conference on pattern recognition; 2010. p. 2366–9.
- [31] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process* 2004;13:600–12.
- [32] Wolterink JM, Leiner T, Viergever MA, Isgum I. Generative adversarial networks for noise reduction in low-dose CT. *IEEE Trans Med Imaging* 2017;36:2536–45.
- [33] Wang J, Noble JH, Dawant BM. Metal artifact reduction for the segmentation of the intra cochlear anatomy in CT images of the ear with 3D-conditional GANs. *Med Image Anal* 2019;58:101553. <https://doi.org/10.1016/j.media.2019.101553>.